Appendix

Appendix A1.1   Study characteristics: Long, Gueron, Wood, Fisher, & Fellerath, 1996 (randomized controlled trial)

Characteristic	Description
Study citation	Long, D., Gueron, J. M., Wood, R. G., Fisher, R., & Fellerath, V. (1996). *LEAP: Three-Year Impacts of Ohio's Welfare Initiative to Improve School Attendance among Teenage Parents.* New York: Manpower Demonstration Research Corporation.
Participants	Between 1989 and 1991 a total of 7,017 pregnant women and custodial parents under 20 years old who were receiving Aid to Families with Dependent Children (AFDC) and did not have a high school diploma or a GED were randomly assigned to the LEAP evaluation—80% to the intervention group and 20% to the control group. Due to changes in the program structure and implementation, 4,050 teens were excluded from the analysis. First, 1,442 older teens who experienced LEAP only during its start-up phase, when the program was undergoing problems operating under rules and procedures different from those that were eventually adopted, were excluded. Second, 2,608 teens who were randomly assigned during the first year of program operations were also excluded so that the analysis sample was limited to teens who were enrolled after the bonus and sanction process was functioning smoothly. The remaining 2,967 teens were chosen as the research sample because they were more representative of participants in an ongoing LEAP program. Because exclusions were based on age at random assignment or date of random assignment, treatment-control equivalence was not disrupted. A survey was fielded to 1,178 of the remaining teens (all control group members and a random sample of 25% of the intervention group members in seven counties) three years after random assignment. The analysis was conducted on the 913 respondents (446 intervention and 467 control) to this survey. More than 50% of the 913 teens in the three-year survey sample began LEAP when they were 17 or 18 years old, 33% entered the sample when they were 16 or younger, and 11% percent were 19. The average age was just over 17.5. Most participants entered LEAP with one child (71%) or were pregnant with their first child (21%), while few (8%) had two or more children. Just over half (58%) of the survey teens reported that they were enrolled in a junior high, high school, or GED program when they entered LEAP, while 42% were out of school. The average highest grade completed was 9.54. Nearly all teens in the survey sample are female (99%) and had never been married (94%) when they were randomly assigned. More than half (54%) headed their own welfare cases at the time of random assignment, with 40% on a parent's AFDC case and 6% on another's AFDC case. The sample was 67% African-American, 31% white, 2% Hispanic, and 1% other.
Setting	The survey was fielded to teens in seven counties in Ohio: Cuyahoga (Cleveland), Franklin (Columbus), Hamilton (Cincinnati), Lawrence, Lucas (Toledo), Muskingum, and Stark.
Intervention	The program had a three-tiered incentive structure: grant increases ($62 for proof of enrollment plus $62 for each month in which they met attendance requirements), grant reductions ($62 for each month they failed to attend an initial assessment interview, failed to verify enrollment in school, or exceeded the allowed number of excused absences), and unchanged grants (exceeded the allowed number of total absences but not the allowed number of unexcused absences). Teens' enrollment and attendance were monitored by case managers, who also offered guidance and authorized assistance with child care and transportation for teens complying with the rules. The $200 bonus for school completion was not provided during the period of the study.
Comparison	Teens in the control group received normal cash benefits, with no bonuses paid or sanctions imposed for school enrollment and attendance.
Primary outcomes and measurement	Outcomes in each of the domains are included in this study: the measure of staying in school is defined as ever or currently enrolled in high school or a GED program; the measure related to progressing in school is ever completing grade 11, because it is the latest measure of progress; and the completion indicator is a combined measure of ever graduated high school or received a GED. All measures are taken from the survey fielded three years after random assignment.
Teacher training	Information on staff training was not available.

Appendix A1.2   Study characteristics: Mauldon, Malvin, Stiles, Nicosia, & Seto, 2000 (randomized controlled trial with attrition problems)

Characteristic	Description
Study citation	Mauldon, J., Malvin, J., Stiles, J., Nicosia, N., & Seto, E. (2000). *Impact of California's Cal-Learn Demonstration Project: Final Report.* Berkeley, CA: University of California, UC DATA.
Participants	Between 1994 and 1997 custodial parents and pregnant teens under age 19 on welfare who did not have a high school diploma or GED were randomly assigned using the last two digits of their Social Security number to one of four groups: full Cal-Learn (including case management and financial incentives), case management only, financial incentives only, and no treatment. This WWC report focuses on the financial incentives and no treatment conditions. The evaluation samples were selected from the lists of all teens in each research county who appeared to be Cal-Learn eligible based on electronic and paper records, which indicated that they were pregnant or custodial teen parents on welfare. After males and those registered in error were removed, the sample was 4,859, and 2,682 of those teens responded to the Wave I survey. Interviews for the Wave I survey were conducted between April 1996 and April 1999, with an average of 13 months between program entry and interview. Additional exclusions were made for teens who lost custody of their children, moved to a nonresearch county or out of state, left AFDC, or did not participate for at least six months. After the additional exclusions, the survey evaluation sample consisted of 2,156 respondents, including 554 in the financial incentives group and 549 in the no treatment group. The Wave II survey was administered for the 2,156 Wave I respondents who were not excluded from the sample, with 1,562 respondents. This data were collected 26 months after program entry, on average, and outcomes from the Wave I survey were used for teens who did not respond to the Wave II survey. Since the study does not present overall findings for the entire sample, this review presents the findings for teens age 18 and older at the most recent survey, which comprises nearly 83% of study teens. For 65% of the teens in the survey sample, Cal-Learn began when they were 17 or 18 years old, while the remaining 35% entered the sample when they were 16 or younger. The average age was 17.2. Most participants (72%) entered Cal-Learn with only one child, while 24% had no children and 4% entered with two or more children. Of teens age 18 or older at their latest interview, 67% reported that they were enrolled in school when they entered Cal-Learn, compared with 33% who were not. All teens in the analysis sample are female. The sample was 47% Hispanic, 27% African-American, 21% white, and 5% other.
Setting	The program was administered by welfare offices in four counties in California: Alameda, Los Angeles, San Bernadino, and San Joaquin.
Intervention	The program featured two key elements: financial bonuses and penalties for either making progress in school or failing to make progress and intensive, individualized case management to help each client move toward high school graduation or GED receipt.[1] Cal-Learn clients who graduated with a high school diploma or earned a GED received a $500 reward through a personal check. Prior to graduation, good progress was rewarded with a $100 bonus check applied to the family welfare grant of a Cal-Learn teen when she received a report card indicating satisfactory progress (at least a C average). But if the Cal-Learn teen did not turn in a report card or had a report card showing inadequate progress (a D– or F average), the family welfare check was reduced by a total of $100 (two $50 deductions over two consecutive months). Teens who earned an average grade between C and D were rated as making adequate progress and received neither a bonus nor a sanction. Up to four report cards a year were assessed to determine if bonuses or sanctions were warranted. In addition, all participants who were attending school were entitled to receive subsidies for support services—child care, transportation, and other school-related expenses. Generally, bonuses were issued within a month, while sanctions took two months.
Comparison	Control group students were neither directed toward case management nor eligible to receive bonuses or sanctions. All teens in the evaluation were offered support services, including reimbursement for child care, transportation to school, and school-related expenses.
Primary outcomes and measurement	The study included a measure of staying in school (dropped out) and a measure of completing school (received high school diploma or GED). Dropping out and high school diploma receipt were taken from survey responses, while GED receipt came from administrative records.
Teacher training	Information on staff training was not available.

1. Since this review focuses on the effects of financial incentives for teen parents, the results presented here are for the intervention condition with financial incentives only. A comparison of the financial incentives, case management, and combined treatments is presented in Appendices A4.5 and A4.6.

Appendix A2.1   Outcome measures in the staying in school domain

Outcome measure	Description
Dropped out	This binary measure of whether a student has dropped out of school is based on not having a high school diploma or GED and not being enrolled in high school or a GED program. For LEAP (Long et al., 1996), the measure was taken as the converse of "ever completed high school or GED, or currently enrolled in high school or GED program." The combined measure presented in the report was based on responses to a survey three years after random assignment. For Cal-Learn (Mauldon et al., 2000), teens were counted as having dropped out of high school if, at the time of the latest Retrospective Survey interview, they did not have a GED or a high school diploma and were neither enrolled in school nor a GED program nor on summer break.[1]

1. Surveys were administered, on average, 13 and 26 months after random assignment. The measure was based on responses to the latest survey, approximately 72% from the Wave II survey and 28% from the Wave I survey.

Appendix A2.2   Outcome measures in the progressing in school domain

Outcome measure	Description
Completed 11th grade	This binary measure of completion of grade 11 was obtained from survey data. For LEAP (Long et al., 1996), the measure was based on responses to a survey three years after random assignment.

Appendix A2.3   Outcome measures in the completing school domain

Outcome measure	Description
Completed high school or GED	This binary measure of whether a student has graduated from high school or received a GED is based on survey and school records data. For LEAP (Long et al., 1996), the measure was based on responses to a survey three years after random assignment. For Cal-Learn (Mauldon et al., 2000), high school diploma receipt was taken from survey responses (see footnote 1 in Appendix A2.1). Information on GED receipt was based on data from the California Department of Education. GED data were not available for individual teens but were calculated for each of the subsets of evaluation samples, so it was still possible to compute impacts on GED completion within groups of teens.

Appendix A3.1   Summary of study findings included in the rating for the staying in school domain[1]

| Outcome measure | Study sample | Sample size (students) | Author's findings from the study | | | WWC calculations | | |
| | | | Mean outcome | | | | | |
			Financial Incentives group	Comparison group	Mean difference[2]	Effect size[3]	Statistical significance[4] (at $\alpha = 0.05$)	Improvement index[5]
Long et al., 1996 (randomized controlled trial)								
Dropped out	Full sample	913	48.4	53.5	5.1	0.12	ns	+5
Average[6] for staying in school (Long et al., 1996)						0.12	ns	+5
Mauldon et al., 2000 (randomized controlled trial with attrition problems)								
Dropped out	18 and older at survey	906	44.7	52.3	7.6	0.19	Statistically significant	+7
Average[6] for staying in school (Mauldon et al., 2000)						0.19	Statistically significant	+7
Domain average[6] for staying in school						0.16	na	+6

ns = not statistically significant
na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the improvement index. Subgroup findings from the same studies are not included in these ratings, but are reported in Appendices A4.1 and A4.5.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Long et al. (1996) reported rates for "ever completed high school or a GED, or currently enrolled in high school or a GED program." The WWC reports the rates for the members of the sample who are not in this group, that is, dropped out = 1 − [ever finished + still enrolled] = 1 − "ever completed HS / GED or currently in HS / GED." Means from Long et al. (1996) are regression adjusted.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.
6. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect size.

Appendix A3.2   Summary of study findings included in the rating for the progressing in school domain[1]

| Outcome measure | Study sample | Sample size (students) | Author's findings from the study | | | WWC calculations | | |
| | | | Mean outcome | | | | | |
			Financial Incentives group	Comparison group	Mean difference[2]	Effect size[3]	Statistical significance[4] (at $\alpha = 0.05$)	Improvement index[5]
Long et al., 1996 (randomized controlled trial)								
Completed 11th grade	Full sample	913	50.0	45.4	4.6	0.11	ns	+4
Domain average[6] for progressing in school						0.11	ns	+4

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the improvement index. Subgroup findings from the same studies are not included in these ratings, but are reported in Appendix A4.2.

2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Long et al. (1996) are regression adjusted.

3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.

4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.

5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.

6. This row provides the study average, which in this instance is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A3.3   Summary of study findings included in the rating for the completing school domain[1]

| | | | Author's findings from the study | | | WWC calculations | | |
| | | | Mean outcome | | | | | |
Outcome measure	Study sample	Sample size (students)	Financial Incentives group	Comparison group	Mean difference[2]	Effect size[3]	Statistical significance[4] (at α = 0.05)	Improvement index[5]
Long et al., 1996 (randomized controlled trial)								
Completed high school or GED	Full sample	913	34.0	31.9	2.1	0.06	ns	+2
Average[6] for completing school (Long et al., 1996)						0.06	ns	+2
Mauldon et al., 2000 (randomized controlled trial with attrition problems)								
Completed high school or GED	18 and older at survey	906	29.1	24.2	4.9	0.15	ns	+6
Average[6] for completing school (Mauldon et al., 2000)						0.15	ns	+6
Domain average[6] for completing school						0.11	na	+4

ns = not statistically significant

na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the improvement index. Subgroup findings from the same studies are not included in these ratings, but are reported in Appendices A4.3, A4.4, and A4.6.

2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Long et al. (1996) are regression adjusted.

3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.

4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.

5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.

6. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect size.

Appendix A4.1   Summary of subgroup findings for the staying in school domain[1]

| Outcome measure | Study sample | Sample size (students) | Author's findings from the study | | | WWC calculations | | | |
| | | | Mean outcome | | | | | | |
			Financial Incentives group	Comparison group	Mean difference[2]	Effect size[3]	Statistical significance[4] (at α = 0.05)	Improvement index[5]
Long et al., 1996 (randomized controlled trial)								
Dropped out	Initially enrolled[6]	527	34.1	43.1	9.0	0.23	Statistically significant	+9
Dropped out	Not initially enrolled[7]	386	67.8	68.4	0.6	0.02	ns	+1
Mauldon et al., 2000 (randomized controlled trial with attrition problems)								
Dropped out	Initially enrolled	536	30.0	39.8	9.8	0.26	ns	+10
Dropped out	Not initially enrolled	288	71.1	68.6	−2.5	−0.07	ns	−3

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in the staying in school domain. The full sample outcomes were used for rating purposes and are presented in Appendix A3.1.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Long et al. (1996) are regression adjusted.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.
6. Initially enrolled refers to students who were enrolled in school or a GED program at the time of random assignment.
7. Not initially enrolled refers to students who were neither enrolled in school nor a GED program at the time of random assignment.

Appendix A4.2   Summary of subgroup findings for the progressing in school domain[1]

| Outcome measure | Study sample | Sample size (students) | Mean outcome | | WWC calculations | | | |
			Financial Incentives group	Comparison group	Mean difference[2]	Effect size[3]	Statistical significance[4] (at α = 0.05)	Improvement index[5]
				Author's findings from the study				
Long et al., 1996 (randomized controlled trial)								
Completed 11th grade	Initially enrolled	527	60.6	58.1	2.5	0.06	ns	+3
Completed 11th grade	Not initially enrolled	386	35.8	28.0	7.8	0.22	ns	+9

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in the progressing in school domain. The full sample outcomes were used for rating purposes and are presented in Appendix A3.2.

2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Long et al. (1996) are regression adjusted.

3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.

4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.

5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.

Appendix A4.3   Summary of initially enrolled subgroup findings for the completing school domain[1]

| | | | Author's findings from the study | | WWC calculations | | | |
| | | | Mean outcome | | | | | |
Outcome measure	Study sample	Sample size (students)	Financial Incentives group	Comparison group	Mean difference[2]	Effect size[3]	Statistical significance[4] (at α = 0.05)	Improvement index[5]
Long et al., 1996 (randomized controlled trial)								
Completed high school or GED	Initially enrolled	527	45.6	38.6	7.0	0.17	ns	+7
Completed high school or GED	Not initially enrolled	386	18.6	22.1	−3.5	−0.13	ns	−5
Mauldon et al., 2000 (randomized controlled trial with attrition problems)								
Completed high school or GED	Initially enrolled	536	38.5	35.0	3.5	0.09	ns	+4
Completed high school or GED	Not initially enrolled	288	14.8	10.5	4.3	0.24	Statistically significant	+9

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in the completing school domain. The full sample outcomes were used for rating purposes and are presented in Appendix A3.3.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Long et al. (1996) are regression adjusted.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.

Appendix A4.4   Summary of item-level findings for the completing school domain[1]

| | | | Author's findings from the study | | WWC calculations | | | |
| | | | Mean outcome | | | | | |
Outcome measure	Study sample	Sample size (students)	Financial Incentives group	Comparison group	Mean difference[2]	Effect size[3]	Statistical significance[4] (at $\alpha = 0.05$)	Improvement index[5]
Long et al., 1996 (randomized controlled trial)								
High school diploma	Full sample	913	22.9	23.5	−0.6	−0.02	ns	−1
High school diploma	Initially enrolled	527	35.6	34.2	1.4	0.04	ns	+2
High school diploma	Not initially enrolled	386	6.7	7.8	−1.1	−0.10	ns	−4
GED receipt	Full sample	913	11.1	8.4	2.7	0.19	ns	+7
GED receipt	Initially enrolled	527	10.0	4.4	5.6	0.53	Statistically significant	+20
GED receipt	Not initially enrolled	386	12.0	14.3	−2.3	−0.12	ns	−5
Mauldon et al., 2000 (randomized controlled trial with attrition problems)								
High school diploma	18 and older at survey[6]	906	19.2	18.1	1.1	0.04	ns	+2
High school diploma	Initially enrolled	536	29.6	28.9	0.7	0.02	ns	+1
High school diploma	Not initially enrolled	288	4.4	3.9	0.5	0.08	ns	+3
GED receipt	18 and older at survey	906	9.8	6.1	3.7	0.31	Statistically significant	+12
GED receipt	Initially enrolled	536	8.9	6.0	2.9	0.26	ns	+10
GED receipt	Not initially enrolled	288	10.4	6.5	3.9	0.31	ns	+12

ns = not statistically significant

1. This appendix presents item-level findings for measures that fall in the completing school domain. The full sample outcomes were used for rating purposes and are presented in Appendix A3.3.
2. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. Means from Long et al. (1996) are regression adjusted.
3. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
4. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.
5. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.
6. The sample sizes for initially enrolled and not initially enrolled do not sum to the total sample of 18 and older at survey, because teens whose school enrollment at Cal-Learn entry could not be ascertained were excluded.

Appendix A4.5   Summary of alternate treatment findings for the staying in school domain[1]

| Outcome measure | Study sample[2] | Sample size (students) | Author's findings from the study | | WWC calculations | | | |
| | | | Mean outcome | | | | | |
			Intervention group[3]	Comparison group	Mean difference[4]	Effect size[5]	Statistical significance[6] (at $\alpha = 0.05$)	Improvement index[7]
Mauldon et al., 2000 (randomized controlled trial with attrition problems)								
Dropped out	Financial incentives and case management	916	42.5	52.3	9.8	0.24	Statistically significant	+9
Dropped out	Financial incentives	906	44.7	52.3	7.6	0.19	Statistically significant	+7
Dropped out	Case management	878	46.5	52.3	5.8	0.14	ns	+6

ns = not statistically significant

1. As noted in Appendix A1.2, the Cal-Learn study included three different treatment conditions along with a control group: financial incentives only, case management only, and a combination of financial incentives and case management. This appendix presents findings from alternative specifications of the treatment for measures that fall in the staying in school domain. The full sample outcomes were used for rating purposes and are presented in Appendix A3.2.

2. In this appendix, study sample refers to the treatment condition being evaluated.

3. Earlier appendices examine the financial incentives intervention only, so the mean of the intervention group is labeled as the "Financial Incentives group." Since one of the treatment conditions presented in this appendix does not include financial incentives, the label has been changed to the more general "Intervention group."

4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.

5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.

6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.

7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.

Appendix A4.6   Summary of alternate treatment findings for the completing school domain[1]

| | | | Author's findings from the study | | | WWC calculations | | |
| | | | Mean outcome | | | | | |
Outcome measure	Study sample[2]	Sample size (students)	Intervention group[3]	Comparison group	Mean difference[4]	Effect size[5]	Statistical significance[6] (at α = 0.05)	Improvement index[7]
Mauldon et al., 2000 (randomized controlled trial with attrition problems)								
Completed high school or GED	Financial incentives and case management	916	31.5	24.2	7.3	0.22	Statistically significant	+9
Completed high school or GED	Financial incentives	906	29.1	24.2	4.9	0.15	ns	+6
Completed high school or GED	Case management	878	27.0	24.2	2.8	0.09	ns	+4

ns = not statistically significant

1. As noted in Appendix A1.2, the Cal-Learn study included three different treatment conditions along with a control group: financial incentives only, case management only, and a combination of financial incentives and case management. This appendix presents findings from alternative specifications of the treatment for measures that fall in the completing school domain. The full sample outcomes were used for rating purposes and are presented in Appendix A3.3.
2. In this appendix, study sample refers to the treatment condition being evaluated.
3. Earlier appendices examine the financial incentives intervention only, so the mean of the intervention group is labeled as the "Financial Incentives group." Since one of the treatment conditions presented in this appendix does not include financial incentives, the label has been changed to the more general "Intervention group."
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The level of statistical significance was reported by the study authors.
7. The improvement index represents the difference between the percentile rank of the average student in the intervention condition and that of the average student in the comparison condition. The improvement index can take on values between −50 and +50, with positive numbers denoting favorable results.

Appendix A5.1 *Financial Incentives for Teen Parents* rating for the staying in school domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of staying in school, the WWC rated *Financial Incentives for Teen Parents* as having potentially positive effects. It did not meet the criteria for positive effects because it had only one study that showed statistically significant positive outcomes in this domain. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Financial Incentives for Teen Parents* was assigned a potentially positive rating.

Rating received

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

   Met. One study of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important positive effect in this domain.

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

   Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important negative effects in this domain, while one study showed an indeterminate effect in this domain.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

   Not met. Only one study of *Financial Incentives for Teen Parents* showed a statistically significant positive effect in this domain.

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

   Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important negative effects in this domain.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain level effect. The WWC also considers the size of the domain level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Financial Incentives for Teen Parents rating for the progressing in school domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of progressing in school, the WWC rated *Financial Incentives for Teen Parents* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects because it had only one study, and that study showed no statistically significant outcomes, either positive or negative, in this domain.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative.*

    Met. The one study of *Financial Incentives for Teen Parents* showed no statistically significant or substantively important effect in this domain.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

    Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant positive effect in this domain.

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

    Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important negative effects in this domain.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

    Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important positive effect in this domain.

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

    Not met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important negative effects, while one study showed an indeterminate effect in this domain.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

    Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important effect in this domain.

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

    Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important effect in this domain.

(continued)

Other ratings considered

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

    Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important negative effect in this domain.

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

    Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important positive effects in this domain.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a strong design.

    Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant negative effect in this domain.

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

    Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important positive effects in this domain.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain level effect. The WWC also considers the size of the domain level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5.3   *Financial Incentives for Teen Parents* rating for the completing school domain

The WWC rates an intervention's effects for a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of completing school, the WWC rated *Financial Incentives for Teen Parents* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects because neither of the two studies showed statistically significant outcomes, either positive or negative, in this domain.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative.*

  Met. The two studies of *Financial Incentives for Teen Parents* showed no statistically significant or substantively important effects in this domain.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant positive effect in this domain.

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

  Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important negative effects in this domain.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

  Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important positive effect in this domain.

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

  Not met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important negative effects in this domain, while two studies showed indeterminate effects in this domain.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

  Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important effect in this domain.

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

  Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important effect in this domain.

(continued)

Other ratings considered

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

  Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant or substantively important negative effect in this domain.

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

  Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important positive effects in this domain.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. No studies of *Financial Incentives for Teen Parents* showed a statistically significant negative effect in this domain.

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

  Met. No studies of *Financial Incentives for Teen Parents* showed statistically significant or substantively important positive effects in this domain.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain level effect. The WWC also considers the size of the domain level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.